Payer Type and Low-Value Care: Comparing Choosing Wisely Services across Commercial and Medicare Populations

Carrie H. Colla, Nancy E. Morden, Thomas D. Sequist, Alexander J. Mainor, Zhonghe Li, and Meredith B. Rosenthal

Objective. To compare low-value health service use among commercially insured and Medicare populations and explore the influence of payer type on the provision of low-value care.

Data Sources. 2009–2011 national Medicare and commercial insurance administrative data.

Design. We created claims-based algorithms to measure seven Choosing Wisely-identified low-value services and examined the correlation between commercial and Medicare overuse overall and at the regional level. Regression models explored associations between overuse and regional characteristics.

Methods. We created measures of early imaging for back pain, vitamin D screening, cervical cancer screening over age 65, prescription opioid use for migraines, cardiac testing in asymptomatic patients, short-interval repeat bone densitometry (DXA), preoperative cardiac testing for low-risk surgery, and a composite of these.

Principal Findings. Prevalence of four services was similar across the insurance-defined groups. Regional correlation between Medicare and commercial overuse was high (correlation coefficient = 0.540–0.905) for all measures. In both groups, similar region-level factors were associated with low-value care provision, especially total Medicare spending and ratio of specialists to primary care physicians.

Conclusions. Low-value care appears driven by factors unrelated to payer type or anticipated reimbursement. These findings suggest the influence of local practice patterns on care without meaningful discrimination by payer type.

Key Words. Low-value care, overuse, waste, regional variation

Health care spending varies broadly across regions of the United States, but it results in few differences in outcomes (survival, function), patient satisfaction, physician satisfaction, or access to care (Fisher et al. 2003a, 2003b; Sirovich et al. 2006; Fowler et al. 2008; Anthony et al. 2009; Fisher, Bynum, and Skinner 2009).

Researchers estimate 30 percent of U.S. health care spending is wasted (Smith et al. 2013). Addressing this concerning waste, the Choosing Wisely initiative was launched by the American Board of Internal Medicine (ABIM) Foundation in 2012 to identify low-value services, improve health care quality, and reduce overuse. Through the Choosing Wisely campaign, physician societies have begun to identify and revise lists of specialty-specific, low-value services whose avoidance would increase the quality and value of care provided (American Board of Internal Medicine 2015). The next essential step in reducing health services overuse is leveraging these lists to develop interventions effective in reducing low-value services.

An exploration of regional patterns of Choosing Wisely–identified low-value care among Medicare beneficiaries has advanced our understanding of health care variation. This work has demonstrated regional differences not simply in overall health care intensity and spending but in low-value services specifically, bringing us closer to understanding waste and its causes (Fisher et al. 2003a; Baicker et al. 2004; Schwartz et al. 2014; Colla et al. 2015). Nonetheless, determinants of low-value care provision remain incompletely understood.

In addition to practice patterns and population factors, differences in insurance design could substantially influence the provision of low-value health care. In Medicare, reimbursement rates are set by the federal government and applied uniformly (for the same procedure) with geographic adjustments to capture local cost differences. In contrast, commercial insurers negotiate prices directly with providers, allowing for substantially more variability in reimbursements, and typically higher reimbursement rates than Medicare (Medicare Payment Advisory Commission 2011). For most providers and services, profit margins for delivering care to commercially insured patients are higher and thus incentives to recommend services stronger, all else (such as need) being equal. By examining geographic variation in low-value care through the lens of these two distinct payment mechanisms, we can explore the relative influence of reimbursement and underlying practice patterns on region-level low-value care provision.

———————

Address correspondence to Carrie H. Colla, Ph.D., The Dartmouth Institute, Level 5 WTRB, One Medical Center Drive, Lebanon, NH 03756; e-mail: carrie.h.colla@dartmouth.edu. Nancy E. Morden, M.D., M.P.H., and Alexander J. Mainor, J.D., M.P.H., are with the The Dartmouth Institute, Lebanon, NH. Thomas D. Sequist, M.D., M.P.H., is with the Department of Health Care Policy, Harvard Medical School, Boston, MA; Partners Healthcare System, Boston, MA; and Brigham and Women's Hospital Division of General Medicine and Primary Care, Boston, MA. Zhonghe Li, M.S., is with the Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, MA. Meredith B. Rosenthal, Ph.D., is with the Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, MA.

In this study, we explored the association between payer type and provision of seven Choosing Wisely–identified low-value services. We employed aggregated, multiple commercial payers' claims data from the Health Care Cost Institute (HCCI) as well as Medicare administrative data to examine prevalence and correlates of this low-value care at the national and Hospital Referral Region (HRR) level for both Medicare and commercially insured populations (The Dartmouth Atlas of Health Care 2012). We hypothesized overuse may be more prevalent in the older Medicare population as this group has more frequent contact with the health care system, but higher reimbursement rates in the commercial population may have the opposite effect.

METHODS

We compared the prevalence of seven Choosing Wisely services between Medicare and commercially insured populations and examined correlations between the two populations for significance. We used standardized z-scores for these services to create a composite measure of overuse across services for the commercial and Medicare populations. We then used multivariate regression analysis to study the association between overuse and HRR-level covariates separately within each database, as we have performed in prior research (Colla et al. 2015).

Medicare and Commercial Claims Data

We used 100 percent Medicare administrative claims data (2009–2011) to determine the prevalence of low-value services in the Medicare population. We limited our analysis to fee-for-service beneficiaries enrolled in Medicare Parts A and B (inpatient and outpatient insurance). We also required enrollment in Medicare Part D (prescription insurance) for two measures of Choosing Wisely services related to prescription drugs; analyses employing Part D data were limited to a 40 percent random sample. Commercial insurance claims data were obtained from the Health Care Cost Institute (HCCI) for 2009–2011. This dataset includes nationally representative commercial data from three of the nation's largest insurers: Aetna, Humana, and UnitedHealthcare. Commercial claims data gathered by HCCI represent the health care activity of more than 50 million individuals per year. Medicare and commercially insured patients were assigned to Dartmouth Atlas hospital referral regions (HRRs) using residential ZIP code.

Choosing Wisely Measurement

We developed claims-based algorithms for seven services, representing 25 Choosing Wisely recommendations (due to common services on multiple specialty society lists) employing previously published methodology (Colla et al. (2015) (see Tables S1, S2). We used the same definitions to define the HCCI and Medicare cohorts. Differences in available data between the Medicare and HCCI cohorts permitted examination of just 7 of the 11 measures previously studied. Measures include early imaging for back pain, vitamin D screening for low-risk patients, cervical cancer screening for patients over age 65, prescription opioid use for migraines, cardiac testing in asymptomatic patients, short-interval dual-energy X-ray absorptiometry (bone density) testing, preoperative cardiac testing in low-risk patients preceding low-risk (noncardiac) surgery, and a composite of these.

Area-Level Variables

Using the methodology in Colla et al. (2015), we employed HRR-level covariates in an exploratory regression analysis. These measures included Medicare price-, age-, sex-, race-adjusted per-beneficiary health care spending (a measure of health services use intensity); physician group concentration (a measure of market competition); the ratio of specialists to primary care physicians; a quality of care score (based on effective care measures in the Medicare population); age-, sex-, and race-adjusted mortality rate; the percent of adults reporting fair or poor health; the percent of Medicare beneficiaries of black race; the percent of Medicare beneficiaries of Hispanic ethnicity; the percent of HRR residents living in a rural area; and the percent of residents below 150 percent of the federal poverty level. The quality of care score is generated by the Dartmouth Atlas and includes five effective care services (beta-blockers among heart attack survivors, mammography among women aged 67–69, glycosylated hemoglobin monitoring among diabetics, and retinal examinations among diabetics); details are available on the Dartmouth Atlas website (The Dartmouth Atlas of Health Care 1999).

Statistical Analysis

For each insurance cohort, we calculated an average annual prevalence of each Choosing Wisely service in the at-risk population, both nationally and at the HRR level. We calculated interquartile range (IQR) across HRRs. We

constructed an overall composite measure of low-value care for each HRR, equal to the average of the z-scores for the seven measures (Colla et al. 2015). The Cronbach's alpha for the resulting composite measure is 0.23 for Medicare patients and 0.49 for commercially insured patients, suggesting weak agreement across the seven measures. We examined geographic variation by dividing HRRs into quintiles of performance on the composite measure and mapping the results, aggregating all years of our data for greater statistical stability (Figures 1 and 2). We used ordinary least squares regression to determine the association of HRR-level characteristics with the composite low-value care scores ($N = 306$ HRRs).

Statistical analyses were performed using *SAS* (Cary, NC, USA) and *STATA* 13 (College Station, TX, USA) software. The study was approved by the institutional review boards at Dartmouth College and the Harvard T.H. Chan School of Public Health.

RESULTS

Across payer types, use of low-value care varied little for nearly all measures (Table 1 and Figure 3). Cardiac screening in low-risk, asymptomatic patients
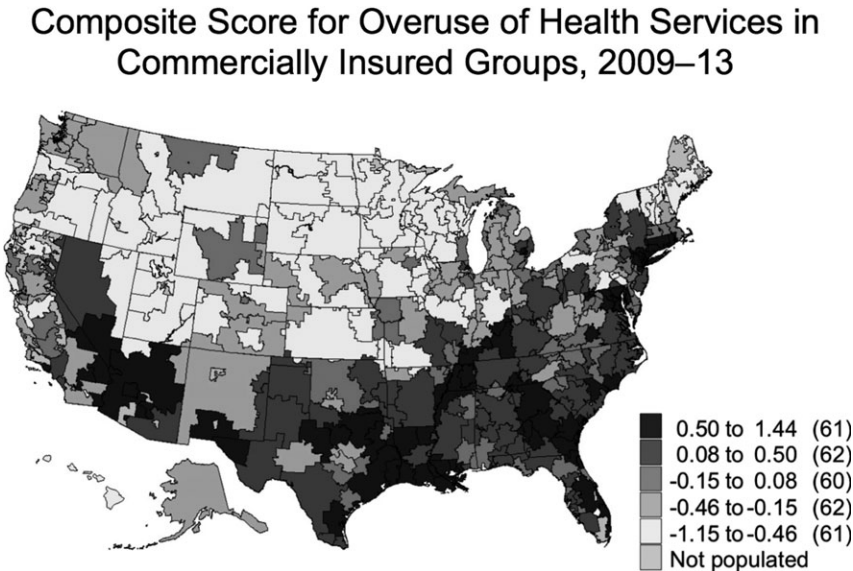
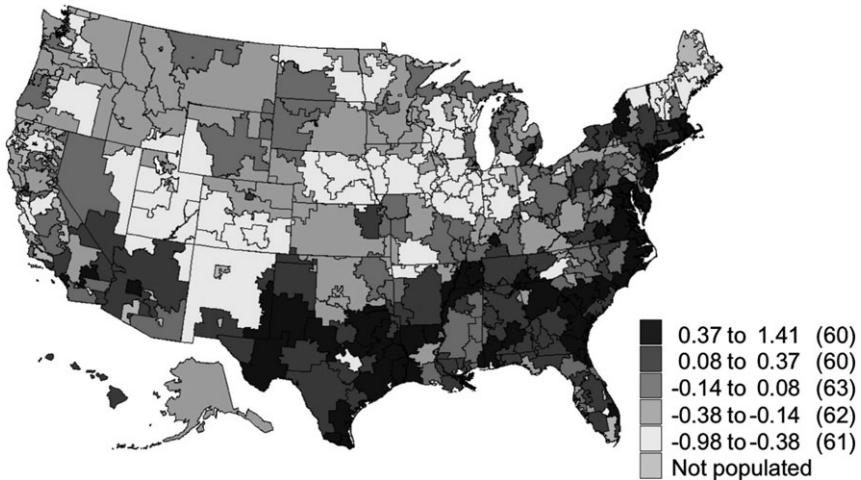Figure 1:   Geographic Variation—Commercial

Composite Score for Overuse of Health Services in Commercially Insured Groups, 2009–13

Figure 2: Geographic Variation—Medicare

Composite Score for Overuse of Health Services in Medicare Insured Groups, 2006–11

■	0.37 to 1.41	(60)
■	0.08 to 0.37	(60)
▨	-0.14 to 0.08	(63)
▨	-0.38 to -0.14	(62)
□	-0.98 to -0.38	(61)
▨	Not populated	

Notes. Hospital Referral Region (HRR)-level analysis provides statistically stable denominator populations for our estimates of service use prevalence. However, aggregating at this level likely obscures differences among health care providers within an HRR. The above maps reflect the different available years of data between the two datasets. Only years 2009–2011 were used in analysis, but the full spread of available data is used here to reduce the presence of unpopulated HRRs.

was statistically equivalent between the two cohorts ($p = .22$; IQR for Medicare: 9.1–12.5 percent, commercial: 6.2–15.9 percent), but the commercial group had a much larger population potentially affected (in part because the measure definition excluded patients over age 80 from the low-risk pool). Three measures were statistically different but clinically similar across both populations: short-interval dual-energy X-ray absorptiometry (DXA) use ($p =< .0001$; IQR for Medicare: 5.5–10.3 percent, commercial 3.4–12.4 percent), use of opioids in migraine patients ($p = .007$; IQR for Medicare: 18.8–29.9 percent, commercial: 17.4–35.9 percent), and cervical cancer screening over age 65 ($p =< .0001$; IQR for Medicare: 6.9–9.6 percent, commercial: 5.0–13.4 percent). Vitamin D screening was slightly more common in the Medicare than commercial population ($p =< .0001$, IQR for Medicare: 8.8–16.1 percent, commercial: 3.9–13.3 percent). Two services had a larger difference across insurance populations than others: back pain imaging was more common among the commercially insured ($p =< .0001$; IQR for Medicare:

Table 1:   Low-Value Services in Medicare and Commercially Insured Populations (Hospital Referral Region Level)

Choosing Wisely Low-Value Service	Medicare					Commercial					Correlation
	Population. Eligible (Millions) Total Pop. = 515.0M	Medicare Population Affected (Millions)	Medicare Mean Prevalence	Medicare IQR	Medicare Min–Max	Population Eligible (Millions) Total Pop. = 418.4M	Commercial Population Affected (Millions)	Commercial Mean Prevalence	Commercial IQR	Commercial Min–Max	
Back pain imaging	8.4	1.9	22.6%	20.3–25.1%	12.8–32.9%	3.1	0.9	28.8%	21.8–35.3%	15.8–46.1%	0.54
Vitamin D screening	159.4	20.4	12.8%	8.8–16.1%	3.6–29.3%	255.6	20.7	8.1%	3.9–13.3%	2.1–18.6%	0.90
Cervical cancer screening	172.3	14.3	8.3%	6.9–9.56%	3.7–17.4%	16.5	1.5	9.1%	5.0–13.4%	2.6–16.3%	0.66
Opioids in migraine patients	0.4	0.1	24.5%	18.8–29.9%	10.7–49.5%	0.4	0.1	26.1%	17.4–35.9%	8.0–57.7%	0.63
Cardiac screening	30.6	3.4	11.1%	9.1–12.5%	6.2–20.7%	133.3	14.4	10.8%	6.2–15.9%	4.0–25.3%	0.82
DXA testing	143.0	12.2	8.5%	5.5–10.3%	2.7–32.5%	8.0	0.6	7.5%	3.4–12.4%	1.2–21.0%	0.80
Preoperative cardiac testing	0.9	0.4	45.9%	39.6–52.6%	15.0–73.9%	1.5	0.4	26.1%	16.4–35.2%	5.3–52.6%	0.77

Notes. "Back pain imaging" is the average annual percent of patients with uncomplicated, incident low-back pain who received nonindicated low-back-pain imaging in the 6 weeks following diagnosis, 2010–2011.

"Vitamin D screening" is the average annual percent of low-risk patients who received at least one nonindicated vitamin D screening test, 2009–2011.

"Cervical cancer screening" is the average annual percent of female patients who received at least one nonindicated screening test for cervical cancer, 2009–2011.

"Opioids in migraine patients" is the average annual percent of patients with a diagnosed migraine who received a nonindicated opioid prescription in the 21 days after an office visit with migraine diagnosis, 2009–2011.

"Cardiac screening" is the average annual percent of low-risk patients who received one or more nonindicated cardiac tests, 2009–2011.

"DXA testing" is the average annual percent of nonindicated dual-energy X-ray absorptiometry (DXA) tests performed within 23 months of a previous DXA test, 2011.

"Preoperative cardiac testing" is the average annual percent of patients undergoing low-risk, noncardiac surgery who received one or more nonindicated cardiac tests in the 30 days before surgery, 2009–2011.

All levels at the person level, with the exception of DXA testing, which is at the test level.

IQR = Interquartile range, 25th to 75th percentile.

Correlation = Pearson correlation coefficient.

This analysis builds on our earlier work in the Medicare population to define and measure overuse of Choosing Wisely low-value services. Two considerations are important when comparing the two analyses, however. First, to compare the Medicare and commercially insured populations, we removed four services from our claims-based Choosing Wisely analysis because commercial data were unavailable or the services were not applicable for those under 65 (Colla et al. 2015). Second, while our initial Medicare analysis used claims data from 2006 to 2011, comparable commercial data for this study were only available from 2009 to 2011. These differences could explain the difference in significance of some Medicare variables in the multivariate regression from our original analyses to this study.

Figure 3: Mean Annual Prevalence of Choosing Wisely Service Use, 2009–2011

* p<0.05, ** p<0.01, *** p<0.001   ■ Medicare ■ Commercial
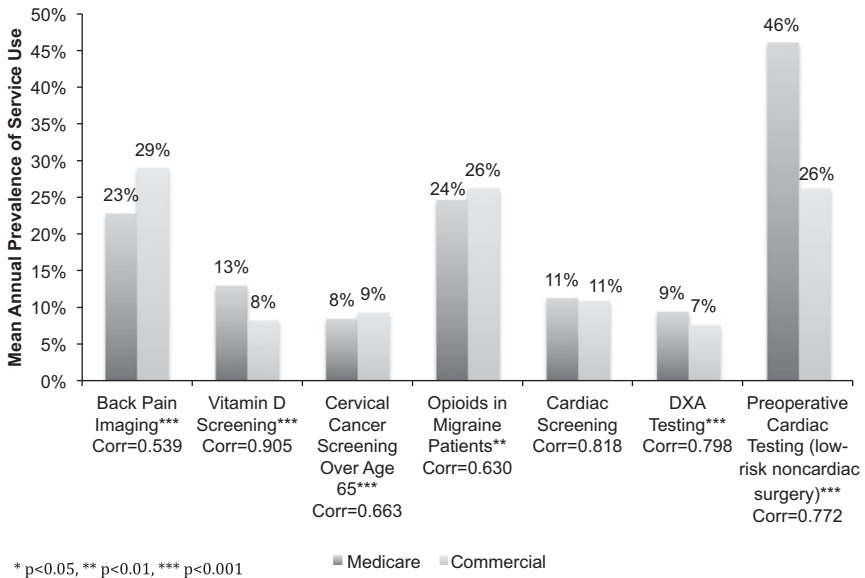
Notes. "Back Pain Imaging" is the average annual percent of beneficiaries with uncomplicated, incident low-back pain who received nonindicated low-back-pain imaging in the 6 weeks following diagnosis, 2010–2011. "Vitamin D Screening" is the average annual percent of low-risk beneficiaries who received at least one nonindicated vitamin D screening test, 2009–2011. "Cervical Cancer Screening" is the average annual percent of female beneficiaries who received at least one nonindicated screening test for cervical cancer, 2009–2011. "Opioids in Migraine Patients" is the average annual percent of beneficiaries with a diagnosed migraine who received a nonindicated opioid prescription in the 21 days after an office visit with migraine diagnosis, 2009–2011. "Cardiac Screening" is the average annual percent of low-risk beneficiaries who received one or more nonindicated cardiac tests, 2009–2011. "DXA Testing (short interval)" is the average annual percent of nonindicated dual-energy X-ray absorptiometry (DXA) tests performed within 23 months of a previous DXA test, 2011. "Preoperative Cardiac Testing (low-risk noncardiac surgery)" is the average annual percent of beneficiaries undergoing low-risk, noncardiac surgery who received one or more nonindicated cardiac tests in the 30 days before surgery, 2009–2011.

20.3–25.1 percent, commercial: 21.8–35.3 percent); preoperative cardiac testing for low-risk noncardiac surgery was more common among Medicare beneficiaries ($p =< .0001$; IQR for Medicare: 39.6–52.6 percent, commercial: 16.4–35.2 percent).

The correlation between the Medicare and commercial rates across HRRs for preoperative cardiac testing was 0.77. The correlation between the groups in low-back pain imaging was 0.54. Use of short-interval bone

densitometry was slightly higher in the Medicare versus commercial population (9 percent vs. 8 percent, respectively) with a correlation between the groups of 0.80. Vitamin D screening, while higher in the Medicare population (13 percent vs. 8 percent in commercial), had the highest correlation between insurance cohorts (0.90).

For most measures, the prevalence of each low-value service was relatively constant over time. However, vitamin D screening increased substantially from 2009 to 2011, from 10 to 16 percent in the Medicare population and from 5 to 10 percent in the commercial group. The only measure of overuse that fell during the study period was cervical cancer screening in those aged 65 and older, declining from 9 to 8 percent in Medicare and 11 to 7 percent in the commercial group from 2009 to 2011.

Some HRR-level health system characteristics were associated with provision of the seven low-value services and the composite measure in both the commercial and Medicare cohorts (Table 2), but most associations, while statistically significant, were relatively weak. The strongest association was observed for higher specialist to primary care physician ratio (Medicare Coeff 0.46 [CI: 0.22–0.71], commercial Coeff 0.58 [CI: 0.37–0.80]). A composite measure of quality that captures use of effective care (e.g., beta-blockers among heart attack survivors, appropriate diabetes care) was positively and significantly related to overuse in the Medicare population (Coeff 0.19 [CI: 0.11–0.27]), but not in the commercial population. Adjusted per capita total Medicare spending was positively associated with overuse in both populations, indicating areas with more health services use overall have more low-value services use (Medicare Coeff 0.10 [CI: 0.05–0.15], commercial Coeff 0.11 [CI: 0.07–0.15]). Finally, physician group concentration was negatively associated with overuse, indicating areas with more physician group competition have more overuse (Medicare Coeff −0.02 [CI: −0.03 to −0.01], commercial Coeff −0.01 [CI: −0.02 to −0.00]).

Hospital Referral Region–level patient population characteristics were also significantly but weakly related to overuse. Percent with poor or fair health living in the area was positively related to overuse in the commercial population (Coeff 0.02 [CI: 0.00–0.04]). Percent black race was positively related to overuse in both groups (Medicare Coeff 0.02 [CI: 0.01–0.02]), commercial Coeff 0.02 [CI: 0.01–0.03]). The local poverty measure (proportion of residents in the HRR below 150 percent of the federal poverty level) was inversely related to overuse in the commercial population (Coeff −0.02 [CI: −0.03 to −0.00]). The mortality rate, the percent Hispanic ethnicity, and the percent living in rural areas were unrelated to overuse.

Table 2: Regional Predictors of Composite Choosing Wisely Service Use Measures, 2009–2011 (*N* = 306 Hospital Referral Regions)

Characteristics	Mean for Characteristic	Medicare Coefficient (95% CI)	Commercial Coefficient (95% CI)
Regional health system characteristics			
Adjusted Medicare reimbursement	$9,570	0.10*** (0.05, 0.15)	0.11*** (0.07, 0.15)
Physician group concentration	0.68	−0.02** (−0.03, −0.01)	−0.01* (−0.02, −0.00)
Specialist/primary care ratio	1.76	0.46** (0.22, 0.71)	0.58*** (0.37, 0.80)
Quality score	0.00	0.19*** (0.11, 0.27)	0.03 (−0.04, 0.10)
Regional population characteristics			
Mortality rate	4.85%	−0.13 (−0.29, 0.03)	0.05 (−0.09, 0.19)
Percent with poor or fair health	15.71%	0.02 (−0.01, 0.04)	0.02* (0.00, 0.04)
Percent black	7.44%	0.02*** (0.01, 0.02)	0.02*** (0.01, 0.03)
Percent Hispanic	4.92%	0.01 (−0.00, 0.01)	0.01 (−0.00, 0.01)
Percent rural	22.06%	0.00 (−0.00, 0.01)	−0.00 (−0.00, 0.00)
Poverty rate	22.88%	−0.01 (−0.03, 0.00)	−0.02** (−0.03, −0.00)

Notes. The dependent variable in this linear regression is an aggregate composite of overuse created by calculating the mean of the z-scores for each of the seven services.

"Adjusted Medicare reimbursement" is the total, average, annual age-, sex-, race- and price-adjusted reimbursement per beneficiary in thousands of dollars (The Dartmouth Atlas of Health Care 2017).

"Physician group concentration" is the Herfindahl–Hirschman Index of allowed Medicare charges to physician groups (the mean of the square of each provider tax identification number's allowed charges divided by total allowed charges within a given HRR) (CMS Research Data Assistance Center 2016).

"Specialist/primary care ratio" is the ratio of specialists per 100,000 residents to primary care physicians per 100,000 residents (The Dartmouth Atlas of Health Care 2017).

"Quality score" is a composite of the standardized rates for the following measures: beneficiaries filling at least one prescription for beta-blockers within 6 months of a heart attack: the percent of female beneficiaries aged 67–69 who received mammography every two years; the percent of diabetics who received appropriate hemoglobin monitoring; and the percent of diabetics who received appropriate eye examinations (The Dartmouth Atlas of Health Care 2012).

"Mortality rate" is age- sex-, and race-adjusted mortality per 1,000 Medicare enrollees (The Dartmouth Atlas of Health Care 2017).

"Percent with poor or fair health" is the percent of adults that reported fair or poor health in the region (Centers for Disease Control and Prevention 2017).

"Percent black" is the percent of Medicare beneficiaries identified as black (RTI International 2010).

"Percent Hispanic" is the percent of Medicare beneficiaries identified as Hispanic (RTI International 2010).

"Percent rural" is the percent of residents in rural areas (U.S. Census Bureau 2010b).

"Poverty rate" is the percent of residents in the region below 150 percent of the federal poverty limit (U.S. Census Bureau 2010a).

*p < .05, **p < .01, ***p < .001.

DISCUSSION

We find a high correlation between overuse in the Medicare and commercially insured populations across HRRs (ranging from 0.54 to 0.90). The tendency to deliver or avoid low-value care appears largely independent of payer type (Medicare or commercial) and patient population attributes. The finding that patients with commercial insurance do not have higher rates of overuse suggests either the presumed difference in anticipated reimbursement (and in turn profit margins) is unimportant to providers or that providers are unable (or unwilling) to discriminate by payer type at the point of care. Our exploratory regression results indicate common HRR-level factors associated with overuse across payer-type defined populations. Variation across HRRs appears primarily associated with market factors; this appears to be a place effect spanning HRRs, leaving open the possibility for multiple health system effects as well. The HRR specialist to primary care physician ratio is significantly and positively associated with overuse. Higher overall Medicare spending is also positively associated with overuse regardless of the payer. Some HRRs may deliver more overuse due to higher physician group competition, either as a direct result of competition (wherein providers offer larger numbers of services to gain market share) or an indirect result (more competition results in more fragmentation, and in turn, redundant service use).

For the majority of the low-value services, use remained relatively consistent over time. However, over our study period, vitamin D screening increased substantially. This may be the result of increased public awareness and promotion of vitamin D deficiency as a medical concern (Holick 2007; Holick and Chen 2008). Additionally, use of cervical cancer screening in the over 65 population decreased substantially; in 2012, the USPSTF formally categorized the test in the over 65 population as a grade "D" service ("There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.") (US Preventive Services Task Force 2012).

Some low-value services demonstrated larger differences across payer groups. We observed greater differences in preoperative cardiac testing (Medicare 45.9 percent; commercial 26.1 percent) and back pain imaging (Medicare 22.6 percent; commercial 28.8 percent) between the two populations than for other low-value services. Differences in preoperative cardiac testing may reflect differences between the payer group populations; the risk of cardiac conditions increases with advancing age. This likely results in higher levels of uncertainty and concern for meaningful missed diagnoses among older

patients, prompting more testing. Higher rates of back pain imaging in the commercial population do not align as clearly with clinical expectations; the prevalence of this service use varies more at the HRR level among the commercially insured (compared to the Medicare population), suggesting a greater role for practice patterns and system-level influences for this potentially high-margin imaging service (Srinivas, Deyo, and Berger 2012).

Two study results in particular raise questions and concerns. Prescription of opioids for migraine patients is similar in both populations (Medicare 24.5 percent; commercial 26.1 percent), but it has a much higher prevalence than other Choosing Wisely–identified services. These patterns parallel national trends, but use of opioids in our study period may not reflect more recent care patterns and slight declines in prescription opioid use responsive to growing concern over opioid overuse (Bao et al. 2016; Patrick et al. 2016). The observed positive association between black race and receipt of low-value care is also concerning, especially in light of literature associating black race with lower achievement of effective care (Nelson, Smedley, and Stith 2002; Kressin and Groeneveld 2015). This suggests a potential for double jeopardy in health services receipt among black Americans; the topic warrants in-depth exploration.

This is the first study to quantify low-value service differences between Medicare and commercially insured populations nationally. Using both Medicare and commercial data sources, we are not only able to determine low-value service use within each payer population but also the association between service use across payer-type defined populations. The similar prevalence of overuse nationally of each service, comparable patterns of regional variation, and high correlation within HRRs between the two populations suggest a place or health system effect, and little influence of payer differences, driving, or deterring overuse. Previous work compared low-value care among Medicaid and commercially insured populations at the state level (Charlesworth et al. 2016) and similarly found no consistent association between insurance type and provision of low-value care. Work by others (Arora and True 2012; Chen et al. 2014) also suggests local norms of care may develop through training and peer interactions, and these norms may have profound cumulative effects on the costs and quality of care from HRR to HRR. Additionally, unmeasured health system factors can have a strong impact on overuse (Keating et al. 2011; McWilliams et al. 2014). The question remains as to how national-, regional-, and system-level efforts can effectively influence these norms and disseminate high-value norms where low-value practice patterns have developed.

The main limitation of our study is its reliance on administrative data to identify and describe use of low-value services. Claims may not provide

enough information or clinical detail to precisely define and identify low-value care. Some cases we measured may not represent low-value care, although we used conservative definitions and exclusion criteria to minimize this misclassification risk. Another limitation is the range of services measured. While we included Choosing Wisely recommendations on lists from multiple specialty societies and services ordered by both specialists and primary care physicians, the range of services measured is still narrow. Use of a composite measure as an indicator of low-value care is useful to summarize a wide range of procedures across HRRs, but it is less actionable and clinically meaningful than measures for individual procedures, and the weak agreement we observed across the seven measures limits the significance of our composite measure. The use of HRRs for analysis of regional variation reflects trade-offs. HRR-level analysis provides statistically stable denominator populations for our estimates of service use prevalence. However, aggregating at this level likely obscures differences among health care providers within an HRR. It also results in less actionable results as the pattern is attributed to a broad collection of providers. Larger datasets or the study of more common services is needed to reliably examine low-value care provision at the level of the individual system or clinician. Our time frame is short, which has implications for measures that require more observation time to appropriately identify prevalent disease or previous testing (for the back pain and bone density scan cohorts). Our time frame also predates the launch of the Choosing Wisely campaign. In such, our study does not examine the impact of this campaign but rather leverages the Choosing Wisely lists of "Five Things Physicians and Patients Should Question" to define low-value services ripe for study. If providers believed these services to be of high value prior to the Choosing Wisely Campaign, our low-value services label may seem unjust. We believe this scenario unlikely as the Choosing Wisely lists include many services long identified as low value by national and international groups (National Institute for Health and Care Excellence 2016; U.S. Preventive Services Task Force 2016). Our commercial data are limited in some geographic areas where the three payers have smaller market shares. Finally, we do not have data on commercial reimbursement rates or a measure of patient service request, so we cannot examine how these factors affect low-value services provision.

By comparing the prevalence and correlation of overuse between Medicare and commercially insured populations, we demonstrated the lack of association between low-value care provision and payer type, and, by proxy, reimbursement level, in a time when fee-for-service was the dominant reimbursement model. Could the payment model reforms being explored

nationally influence practice patterns more than simple differences in reimbursement? Health reforms promoted through the Affordable Care Act, such as Accountable Care Organizations (ACOs), have the potential to reduce the use of low-value care by focusing on quality and cost reduction simultaneously. Private payers have similar contracts and increasingly are holding providers accountable for cost and quality, paralleling Medicare attempts to reduce low-value care. Recent private payer reforms such as the Alternative Quality Contract have begun to show some success in decreasing use of low-value services (Song et al. 2013). Innovative bundled payments, such as the PROMETHEUS payment model and the Geisinger ProvenCare model, also aim to reduce costs and improve quality (de Brantes, Rosenthal, and Painter 2009; Paulus 2009). Continued research on interventions to reduce low-value care should emphasize potential to scale and spread. Additionally, future research should examine similar data during a time period after the implementation of the Choosing Wisely campaign to determine the campaign's effect, if any, on these baseline cross-sectional results.

Given the challenges of translating Choosing Wisely recommendations into claims-based measures or electronic health record data (Shetty et al. 2015), global payment and shared savings approaches may represent the most promising approach to waste reduction: delegation of decisions about which services are truly low value to front-line providers. Use of system-level interventions, such as defaulting to high-value care in the electronic health record, and use of common, evidence-based protocols, can empower front-line providers to choose appropriate care (Goitein and James 2016). Provider systems are likely better equipped than payers to identify locally acceptable versions of the Choosing Wisely recommendations and leverage culture and collaboration in addition to incentives to achieve meaningful practice changes. Emerging data on both the immediate and longer run effects of delivery system and payment reform will help inform policy makers and researchers about these theories and better target efforts to reduce low-value services.

ACKNOWLEDGMENTS

Care Cost Institute (HCCI) and its data contributors, Aetna, Humana, and UnitedHealthcare, in providing the claims data analyzed in this study.

REFERENCES

American Board of Internal Medicine. 2015. "Lists" [accessed on July 2, 2015]. Available at http://www.choosingwisely.org/doctor-patient-lists/

Anthony, D. L., M. B. Herndon, P. M. Gallagher, A. E. Barnato, J. P. Bynum, D. J. Gottlieb, E. S. Fisher, and J. S. Skinner. 2009. "How Much Do Patients' Preferences Contribute to Resource Use?" *Health Affairs (Millwood)* 28 (3): 864–73.

Arora, A., and A. True. 2012. "What Kind of Physician Will You Be? Variation in Health Care and Its Importance for Residency Training" [accessed on July 7, 2012]. Available at http://www.dartmouthatlas.org/downloads/reports/Residency_report_103012.pdf

Baicker, K., A. Chandra, J. S. Skinner, and J. E. Wennberg. 2004. "Who You Are and Where You Live: How Race and Geography Affect the Treatment of Medicare Beneficiaries." *Health Affairs (Millwood)* Suppl Variation: Var33–44.

Bao, Y., Y. Pan, A. Taylor, S. Radakrishnan, F. Luo, H. A. Pincus, and B. R. Schackman. 2016. "Prescription Drug Monitoring Programs Are Associated with Sustained Reductions in Opioid Prescribing by Physicians." *Health Affairs* 35 (6): 1045–51.

de Brantes, F., M. B. Rosenthal, and M. Painter. 2009. "Building a Bridge from Fragmentation to Accountability — The Prometheus Payment Model." *New England Journal of Medicine* 361 (11): 1033–6.

Charlesworth, C. J., T. H. Meath, A. L. Schwartz, and K. J. McConnell. 2016. "Comparison of Low-Value Care in Medicaid vs Commercially Insured Populations." *JAMA Internal Medicine* 176 (7): 998–1004.

Chen, C., S. Petterson, R. Phillips, A. Bazemore, and F. Mullan. 2014. "Spending Patterns in Region of Residency Training and Subsequent Expenditures for Care Provided by Practicing Physicians for Medicare Beneficiaries." *JAMA* 312 (22): 2385–93.

Colla, C. H., N. E. Morden, T. D. Sequist, W. L. Schpero, and M. B. Rosenthal. 2015. "Choosing Wisely: Prevalence and Correlates of Low-Value Health Care Services in the United States." *Journal of General Internal Medicine* 30 (2): 221–8.

The Dartmouth Atlas of Health Care. 1999. "Appendix on the Geography of Health Care in the United States" [accessed on August 10, 2016]. Available at http://www.dartmouthatlas.org/downloads/methods/geogappdx.pdf

The Dartmouth Atlas of Health Care. 2012. "Quality/Effective Care" [accessed on July 7, 2015]. Available at http://www.dartmouthatlas.org/data/map.aspx?ind=160&ch=35%2c183&tf=34&loct=3&extent=-14071323.4105904872305693.8872850095-7398676.5894095136806306.112714991

The Dartmouth Atlas of Health Care. 2017. "Downloads" [accessed on January 18, 2017]. Available at http://www.dartmouthatlas.org/tools/downloads.aspx

Fisher, E. S., J. P. Bynum, and J. S. Skinner. 2009. "Slowing the Growth of Health Care Costs–Lessons from Regional Variation." *New England Journal of Medicine* 360 (9): 849–52.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas, and E. L. Pinder. 2003a. "The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care." *Annals of Internal Medicine* 138 (4): 273–87.

———. 2003b. "The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care." *Annals of Internal Medicine* 138 (4): 288–98.

Fowler Jr, F. J., P. M. Gallagher, D. L. Anthony, K. Larsen, and J. S. Skinner. 2008. "Relationship Between Regional per Capita Medicare Expenditures and Patient Perceptions of Quality of Care." *JAMA* 299 (20): 2406–12.

Goitein, L., and B. James. 2016. "Standardized Best Practices and Individual Craft-Based Medicine: A Conversation about Quality." *JAMA Internal Medicine* 176 (6): 835–8.

Holick, M. F. 2007. "Vitamin D Deficiency." *New England Journal of Medicine* 357 (3): 266–81.

Holick, M. F., and T. C. Chen. 2008. "Vitamin D Deficiency: A Worldwide Problem with Health Consequences." *American Journal of Clinical Nutrition* 87 (4): 1080s–6s.

Keating, N. L., M. B. Landrum, E. B. Lamont, S. R. Bozeman, S. H. Krasnow, L. N. Shulman, J. R. Brown, C. C. Earle, W. K. Oh, M. Rabin, and B. J. McNeil. 2011. "Quality of Care for Older Patients with Cancer in the Veterans Health Administration Versus the Private Sector: A Cohort Study." *Annals of Internal Medicine* 154 (11): 727–36.

Kressin, N. R., and P. W. Groeneveld. 2015. "Race/Ethnicity and Overuse of Care: A Systematic Review." *Milbank Quarterly* 93 (1): 112–38.

McWilliams, J. M., J. B. Dalton, M. B. Landrum, A. B. Frakt, S. D. Pizer, and N. L. Keating. 2014. "Geographic Variation in Cancer-Related Imaging: Veterans Affairs Health Care System Versus Medicare." *Annals of Internal Medicine* 161 (11): 794–802.

Medicare Payment Advisory Commission. 2011. "Report to the Congress: Medicare and the Health Care Delivery System, Variation in Private-Sector Payment Rates."

National Institute for Health and Care Excellence. 2016. "NICE Savings and Productivity and Local Practice Collections" [accessed on May 4, 2015, 2016]. Available at http://www.nice.org.uk/savingsAndProductivity/collection?page=1&pageSize=2000&type=Donotdo&published=&impact=Unclassified&filter=

Nelson, A. R., B. D. Smedley, and A. Y. Stith. 2002. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care (Full Printed Version)*. Washington, D.C.: National Academies Press.

Patrick, S. W., C. E. Fry, T. F. Jones, and M. B. Buntin. 2016. "Implementation of Prescription Drug Monitoring Programs Associated with Reductions in Opioid-Related Death Rates." *Health Affairs* 35 (7): 1324–32.

Paulus, R. A. 2009. "ProvenCare: Geisinger's Model for Care Transformation through Innovative Clinical Initiatives and Value Creation." *American Health & Drug Benefits* 2 (3): 122–7.

Schwartz, A. L., B. E. Landon, A. G. Elshaug, M. E. Chernew, and J. M. McWilliams. 2014. "Measuring Low-Value Care in Medicare." *JAMA Internal Medicine* 174 (7): 1067–76.

Shetty, K. D., D. Meeker, E. C. Schneider, P. S. Hussey, and C. L. Damberg. 2015. "Evaluating the Feasibility and Utility of Translating Choosing Wisely Recommendations into e-Measures." *Healthcare* 3 (1): 24–37.

Sirovich, B. E., D. J. Gottlieb, H. G. Welch, and E. S. Fisher. 2006. "Regional Variations in Health Care Intensity and Physician Perceptions of Quality of Care." *Annals of Internal Medicine* 144 (9): 641–9.

Smith, M., R. Saunders, L. Stuckhardt, and J. M. McGinnis. 2013. *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America.* Washington, D.C.: National Academies Press.

Song, Z., A. M. Fendrick, D. G. Safran, B. E. Landon, and M. E. Chernew. 2013. "Global Budgets and Technology-Intensive Medical Services." *Healthcare* 1 (1–2): 15–21.

Srinivas, S. V., R. A. Deyo, and Z. D. Berger. 2012. "Application of "Less Is More" to Low Back Pain." *Archives of Internal Medicine* 172 (13): 1016–20.

U.S. Census Bureau. 2010a. "American Community Survey (ACS) Summary File Documentation" [accessed on January 18, 2017]. Available at http://www.census.gov/programs-surveys/acs/technical-documentation/summary-file-documentation.2010.html

U.S. Census Bureau. 2010b. "Census 2010 Datasets" [accessed on January 18, 2017]. Available at http://www2.census.gov/census_2010/

US Preventive Services Task Force. 2012. "Cervical Cancer: Screening" [accessed on May 23, 2012]. Available at http://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/cervical-cancer-screening?ds=1&s=cervicalcancerscreening

U.S. Preventive Services Task Force. 2016. "Recommendations for Primary Care Practice" [accessed on September 1, 2016]. Available at http://www.uspreventiveservicestaskforce.org/Page/Name/recommendations

SUPPORTING INFORMATION

Additional supporting information may be found online in the supporting information tab for this article:

Appendix SA1: Author Matrix.

Table S1: Sample Descriptive Characteristics – Health Care Cost Institute (HCCI).

Table S2: Sample Descriptive Characteristics – Medicare (2009–11).